Claims

What is claimed is:

1.   A method for use in a recommender for evaluating the closeness of two items, each of said items characterized by at least one symbolic feature, said method comprising the steps of:

computing a distance between corresponding symbolic feature values of said two items based on an overall similarity of classification of all instances for each possible value of said symbolic feature values; and

aggregating the distances between each of said symbolic features values to determine the closeness of said two items.

2.   The method of claim 1, wherein said computing step employs a Value Difference Metric (VDM) technique to compute said distance between symbolic features.

3.   The method of claim 1, wherein said computing step employs a modified Value Difference Metric (MVDM) technique to compute said distance between symbolic features.

4.   The method of claim 1, wherein said distance, δ, between two values, V1 and V2, for a specific symbolic feature is given by:

$$\delta(V1, V2) = \sum | C1i/C1 - C2i/C2 |^r$$

wherein C1i is the number of times V1 was classified into class i and C1 is the total number of times V1 occurred in the data set.

5.   The method of claim 1, wherein said items are programs, classes of interest are "watched" and not-watched" and wherein said distance, δ, between two values, V1 and V2, for a specific symbolic feature is given by:

$$\delta(V1, \ V2) = \left| \frac{C1\_watched}{C1\_total} - \frac{C2\_watched}{C2\_total} \right| +$$

$$\left| \frac{C1\_not\_watched}{C1\_total} - \frac{C2\_not\_watched}{C2\_total} \right|$$

wherein C1i is the number of times V1 was classified into class i and C1_total is the total number of times V1 occurred in the data set.

6.   The method of claim 1, wherein one of said items is a cluster mean.

7.   The method of claim 1, wherein said items are programs.

8.   The method of claim 1, wherein said items are content.

9.   The method of claim 1, wherein said items are products.

10.   A method for assigning an item to one or more groups of items, each of said items characterized by at least one symbolic feature, said method comprising the steps of:

computing a distance between corresponding symbolic feature values of said item and at least one item in each of said groups, said distance based on an overall similarity of classification of all instances for each possible value of said symbolic feature values;

aggregating the distances between each of said features values to determine the closeness of said item and at least one item in each of said groups; and

assigning said item to said group associated with a minimum distance value.

11.   The method of claim 10, wherein said computing step employs a Value Difference Metric (VDM) technique to compute said distance between symbolic features.

5   12.   The method of claim 10, wherein said computing step employs a modified Value Difference Metric (MVDM) technique to compute said distance between symbolic features.

13.   The method of claim 10, wherein said distance, δ, 10   between two values, V1 and V2, for a specific symbolic feature is given by:

$$\delta(V1, V2) = \sum | C1i/C1 - C2i/C2 |^r$$

wherein C1i is the number of times V1 was classified into class i and C1 is the total number of times V1 occurred in the data set.

15   14.   The method of claim 10, wherein said items are programs, classes of interest are "watched" and not-watched" and wherein said distance, δ, between two values, V1 and V2, for a specific symbolic feature is given by:

$$\delta(V1, V2) = \left| \frac{C1\_watched}{C1\_total} - \frac{C2\_watched}{C2\_total} \right| +$$

$$\left| \frac{C1\_not\_watched}{C1\_total} - \frac{C2\_not\_watched}{C2\_total} \right|$$

wherein C1i is the number of times V1 was classified into class i and C1_total is the total number of times V1 occurred in the data set.

25

15.   The method of claim 10, wherein one of said items is a cluster mean.

16.   The method of claim 10, wherein said items are 30   programs.

-24-

17.   The method of claim 10, wherein said items are content.

18.   The method of claim 10, wherein said items are products.

19.   A system for use in a recommender for evaluating the closeness of two items, each of said items characterized by at least one symbolic feature, comprising:

a memory for storing computer readable code; and

a processor operatively coupled to said memory, said processor configured to:

compute a distance between corresponding symbolic feature values of said two items based on an overall similarity of classification of all instances for each possible value of said symbolic feature values; and

aggregate the distances between each of said symbolic features values to determine the closeness of said two items.

20.   A system for use in a recommender for evaluating the closeness of two items, each of said items characterized by at least one symbolic feature, comprising:

means for computing a distance between corresponding symbolic feature values of said two items based on an overall similarity of classification of all instances for each possible value of said symbolic feature values; and

means for aggregating the distances between each of said symbolic features values to determine the closeness of said two items.

21.   An article of manufacture for use with a recommender for evaluating the closeness of two items, each of said items characterized by at least one symbolic feature, comprising:

a computer readable medium having computer readable code means embodied thereon, said computer readable program code means comprising:

a step to compute a distance between corresponding symbolic feature values of said two items based on an overall similarity of classification of all instances for each possible value of said symbolic feature values; and

a step to aggregate the distances between each of said symbolic features values to determine the closeness of said two items.

22.   A system for assigning an item to one or more groups of items, each of said items characterized by at least one symbolic feature, comprising:

a memory for storing computer readable code; and

a processor operatively coupled to said memory, said processor configured to:

compute a distance between corresponding symbolic feature values of said item and at least one item in each of said groups, said distance based on an overall similarity of classification of all instances for each possible value of said symbolic feature values;

aggregate the distances between each of said features values to determine the closeness of said item and at least one item in each of said groups; and

assign said item to said group associated with a minimum distance value.

23.   An article of manufacture for assigning an item to one or more groups of items, each of said items characterized by at least one symbolic feature, comprising:

a computer readable medium having computer readable code means embodied thereon, said computer readable program code means comprising:

a step to compute a distance between corresponding symbolic feature values of said item and at least one item in each of said groups, said distance based on an overall similarity of classification of all instances for each possible value of said symbolic feature values;

a step to aggregate the distances between each of said features values to determine the closeness of said item and at least one item in each of said groups; and

a step to assign said item to said group associated with a minimum distance value.